身体活动的双系统理论：一种强化学习的视角*

褚昕宇 [1,3] 王泽军 [2] 肖焕禹 [3]
([1] 上海工程技术大学体育教学部，上海 201620) ([2] 同济大学国际足球学院，上海 200092) ([3] 上海体育学院休闲学院，上海 200438)

摘　要　以理性决策为基础的锻炼行为理论被认为是理解身体活动的主导体系，它提供了与身体活动相关的认知构念作为有价值的信息。基于社会生态模型设计的行为干预措施，因表现出了更好的效果而备受关注。近期研究表明，积极的运动认知和当前体育环境都没能很好地促进个人锻炼习惯的养成，因此有必要探索新的理论体系来阐明个人锻炼习惯的形成机制。解释身体活动的最新体系是双系统理论，由于其考虑了身体活动的无意识和快乐决定因素，有望提供一个更广泛的动机视角。一方面，多个有代表性的身体活动双系统模型，从简单的自发路径，到情境线索与锻炼习惯，再到突出自动情感评价作用的复杂概念模型，阐明了系统 1 的构建，结合锻炼行为理论所关注的系统 2，为模型的构建提供了依据。另一方面，通过对双系统的竞争、协调和层级控制原则的分析，为模型的控制提供了建议。经典的强化学习框架解释了双系统模型的构建与控制原则：在模型的构建方面，无模型与基于模型的强化学习分别表示系统 1 和系统 2。在模型的控制方面，Dyna 协作架构与分层强化学习，为身体活动可能是一种相互协作、分层执行的复杂行动组合提供了合理解释。最后提出强化学习视角下锻炼者-体育环境的互动模式，试图从一个全新的角度探讨锻炼行为。
关键词　身体活动；双系统模型；计划行为理论；锻炼习惯；强化学习

1　引言

经常进行身体活动有助于促进人们的身心健康，因此积极参加体育锻炼，养成良好的个人锻炼习惯十分重要（Rhodes, Janssen, Bredin, Warburton, & Bauman, 2017）。过去的几十年时间里，研究者在理解和预测个人行为方面取得了巨大进展，提出了诸如健康信念模型、计划行为理论、自我决定理论以及跨理论模型等具有深刻影响的锻炼行为理论模型（司琦，2007）。这些理论主要以理性决策为基础，认为人们只要意识到锻炼的益处，就会产生参与锻炼的意向、动机，从而进行有规律的体育锻炼（许昭，毛志雄，2015）。然而迄今为止，依据这些理论设计的行为干预措施都没能有效扩大锻炼者的范围并取得长期而持久的效果（朱为模，2009），积极的运动认知并未发现直接促进儿童青少年活动水平的作用（郭强，汪晓赞，蒋健保，2017）。另外，社会生态模型由于纳入环境因素对个体的影响，它的提出促使对锻炼行为的理论研究从个人水平跃升到多层次的总体水平，并因其表现出更好的干预效果，受到研究者的广泛关注（韩慧，郑家鲲，2016；张展嘉，王正珍，于洪军，陈蔚云，2018）。张加林等人（2017）按日常的生活空间将体育环境划分为家庭、学校、社区及其他体育环境，家庭、学校、社区体育环境又分别位于社会生态模型的人际、组织、社区三个层面上。鉴于环境层面的评分普遍高于行为层面，认为当前体育环境并没能更好地促进儿童青少年锻炼习惯的养成。

笔者推测这可能与两个方面的因素有关：首先，相比较于由知识、态度和技能所组成的个人层面，环境层面的因素作为影响锻炼行为的远端效应，影响力可能相对较弱。其次，多数个人行为是自动化的，受情境线索（context cues）的提示，导致行为几乎不会伴随有

收稿日期：2019-10-24
* 国家社科基金青年项目（17CTY019）资助。
通讯作者：王泽军，E-mail: ddbbt@126.com

意识的思考（Marteau, Hollands, & Fletcher, 2012）。因此有必要探索新的理论体系来阐明个人锻炼习惯的形成机制。用于解释身体活动的最新体系是双系统理论，它认为身体活动是由自动化过程（习惯）和目标定向控制过程（意向）共同调控的复杂行为（Rebar et al., 2016）。从模型的构建与控制两个方面，本研究对已有身体活动相关的双系统模型进行了介绍，并引入强化学习（reinforcement learning, RL）框架，解释双系统模型的构建与控制原则，同时从 RL 角度出发，提出锻炼者-体育环境的互动模式，用以阐释锻炼行为。

2　身体活动的双系统模型

在认知与社会心理学研究中，双系统理论认为行为现象是两种不同的心理过程的结果。1 型加工（系统 1）被认为是快速和自动化的，因为它需要最小的认知资源和努力；2 型加工（系统 2）通常被认为是缓慢和熟虑的，采用更多受控推理（Evans & Stanovich, 2013）。对双系统的描述有许多不同的名称，其中系统 1 被命名为内隐、情境、联想、启发、经验、冲动；对应的系统 2 被命名为外显、抽象、命题、分析、理性、熟虑，这些不同的模型命名通常强调了人的思想、情感和行为的不同方面。虽然理论上习惯的自动化和无意识特征似乎与身体活动所需的努力和复杂性相矛盾，但习惯作为大脑的捷径，允许人们在成功参与日常生活行为的同时，将推理和执行能力留给其他想法和行动。

系统评述与元分析方法的研究表明，身体活动亦属于由目标定向控制过程（意向）和自动化过程（习惯）共同调控的双系统理论范畴（Gardner, Phillips, & Judah, 2011; Rebar et al., 2016）。一方面，对计划行为理论（theory of planned behavior, TPB）应用于身体活动的元分析研究表明，意向与锻炼行为之间的相关性处于中等水平（$r = 0.48$; McEachan, Conner, Taylor, & Lawton, 2011）。另一方面，Rebar 等人（2016）发现，在 37 项关于锻炼习惯的研究中，有 70%的研究显示自我报告测量的习惯与锻炼行为之间存在显著的正相关。并且两篇评述都得出一致的结论，认为习惯与锻炼行为之间存在中等水平的相关性（$r = 0.43$, Gardner et al., 2011; $r = 0.32$, Rebar et al., 2016）。因此，目前的研究证实了意向和习惯分别与锻炼行为之间存在中等强度的正相关，意向与习惯均是影响锻炼行为的重要因素。

尽管以理性决策为基础的锻炼行为理论目前仍是身体活动研究领域中的主导方法，然而考虑到体系的效率问题，研究者也开始强调快速、自动和无意识因素（习惯）在身体活动中的重要作用，并逐步提出和完善了身体活动的双系统模型。笔者认为对于身体活动双系统模型的研究需要涵盖两方面的问题：首先，如何合理构建身体活动双系统模型中系统 1 和系统 2，即模型的构建问题。其次，如何有效控制身体活动双系统模型中系统 1 和系统 2，即模型的控制问题。

2.1　身体活动双系统模型中系统 1 的构建

在系统 1 的构建方面，身体活动整合行为改变（Integrated Behavioral Change, IBC）模型通过整合 TPB、自我决定理论、两阶段模型和双系统模型的主要特点于一体，试图弥补锻炼行为理论模型的不足，但它所构建的自发路径相对简单（Hagger & Chatzisarantis, 2014）。习惯理论模型则强调情境线索是通过自动化过程影响身体活动的，在体育环境与个人锻炼习惯的养成之间构建了一条合理的路径，为改变不良生活方式的干预策略设计提供了新的目标（Wood & Runger, 2016）。而突出自动情感评价（automatic affective evaluations）作用的身体活动双系统模型，通过自动情感评价将情境线索与身体活动联系起来，充分考虑了身体活动的情感因素，可以更好地解释情感启发下的复杂锻炼行为（Conroy & Berry, 2017）。

2.1.1　自发路径

由 Ajzen（1991）提出的 TPB 假设行为意向是影响行为最直接的因素，行为意向反过来受态度、主观规范和主观行为控制的影响（段文婷，江光荣，2008）。TPB 由于其理论模型的简洁性、便于操作性，在锻炼行为改变领域得到了大多数锻炼心理学家的认可（沈梦英 等，2010）。目前，对于如何扩展 TPB 模型以提高它对行为的解释力和预测力，也一度

成为研究者所关注的热点（Ajzen, 2001; Sniehotta, Presseau, & Araujo-Soares, 2014）。近来由 Hagger 和 Chatzisarantis（2014）基于 TPB 提出的身体活动 IBC 模型，描述了来自多种理论的动机与心理构念，对身体活动参与行为的影响的熟虑与自发路径。

其中，熟虑路径是由自我决定理论的自主动机对身体活动的远端效应构建的，这种远端效应是由 TPB 的态度、主观规范、主观行为控制、意向构念来调节的。自发路径包括内隐态度和内隐动机对身体活动的直接影响。Calitri 等人（2009）通过研究发现，当控制由自我报告测量的外显态度时，对身体活动的积极内隐态度与身体活动参与水平呈显著相关。另外，Keatley 等人（2012）的研究主要关注内隐动机对身体活动的影响，提出自主和控制形式的动机可以在外显和内隐两个层面运行，并且内隐动机和外显动机对身体活动均有显著的影响。其中，外显动机对身体活动参与的影响是由意向介导的，而内隐动机的影响是直接的、非中介的，这也证实了内隐动机可以反映自发路径对身体活动的影响。

在 IBC 模型中，TPB 的构念取代了自我决定理论的调节构念，同时该模型也使用了阶段性结构（动机性、意志性），并考虑了熟虑和内隐决定因素对身体活动的作用，弥补了早期锻炼行为理论的不足。然而 IBC 模型仍有一些方面需要改进（Rhodes, 2014）。例如，IBC 模型可能缺乏对身体活动中情感方面的考虑。相比 TPB 中的一般态度构念，自我决定理论的内部调节构念可以更好地构建这个情感领域，但在 IBC 模型中被省略。此外，IBC 模型并没有提及社会生态环境的作用，与熟虑路径相比较，对于自发路径的构建显得太过简单。

2.1.2 情境线索与锻炼习惯

相较于 IBC 模型中简单的自发路径，Wood 和 Runger（2016）提出的习惯理论模型强调情境线索通过激活相应的记忆表征来诱导习惯的应答。随着时间的推移，如果行为发生在具有相似线索的环境中，那么可以假设一种习惯只会引发一种相应的行动提示（cue-to-action），用以替代该行为的熟虑和以目标为基础的决策。从双系统模型的角度看，这种习惯的应答被认为是更有效的默认，因为注意力和努力可以被释放到其他需要关注的方面。只有当系统中出现值得关注的变化时（比如移除的线索，精神状态的改变），才会使注意力回到更有意识和熟虑的系统。

另外，锻炼习惯也适用于上述习惯理论模型。首先，锻炼目标（比如减肥、健身）通过激励人们进行反复的锻炼实践和促使他们接触锻炼情境来影响个人锻炼习惯的形成。一旦形成锻炼习惯，情境线索（比如锻炼场景、其他锻炼者和先前的锻炼体验）会自动激活记忆中的习惯表征（比如锻炼强度、时间与频率）。例如，Neal 等人（2012）发现，当那些具有更强烈跑步习惯的锻炼者接触到与经常跑步的情境相关的词语（比如运动场所名称）时，就会自动联想到跑步与慢跑。并且，对那些定期进行锻炼者的研究发现，大约 90%的人都会有一个特定的锻炼场景或时间线索提示，而那些日常通过特定的场景提示进行锻炼的人，他们对锻炼习惯的应答会更加突出（Tappe & Glanz, 2013）。当经常去校园体育馆的学生们接触到体育馆的场景时，由于习惯了这种场景，他们会主动地提高自己说话的音量（Neal, Wood, Labrecque, & Phillippa, 2012）。

其次，人们不仅会关注到他们已经形成的锻炼习惯，还会关注他们锻炼的目标。除非锻炼者有足够的动机，依据当前的情况调整他们的锻炼行为，否则锻炼习惯就会成为一种默认的应答方式。例如，当锻炼者参加身体活动的意向比平时要弱的时候，他们会倾向于依靠锻炼习惯，只锻炼到习惯的运动强度（Rebar, Elavsky, Maher, Doerksen, & Conroy, 2014）。而深入研究发现，行动规划增加了锻炼习惯相对较弱者的身体活动水平，但与没有行动规划的锻炼者相比，行动规划减少了具有强烈锻炼习惯者的身体活动水平（Maher & Conroy, 2015）。此外，中等强度的习惯会自动受到目标的影响，产生一个习惯强度与现实目标有着互动关系的 U 型曲线（Neal et al., 2012）。许多有跑步习惯的锻炼者强烈认同，他们的跑步活动受到了自身目标的影响。在适当情境下，只有当习惯处于中等强度时才会被目标激活。一旦习惯很强烈，目标几乎没有作用，而无论当前目标如何，情境线索依旧可以激活习惯

的应答。另外，人们会根据他们的锻炼行为来推理自身的锻炼目标。

2.1.3 情境线索、自动情感评价与身体活动

对身体活动的自动情感评价是一种直觉反应，它反映了当锻炼的概念在一个人的头脑中被激活时，迅速、非自愿产生的情感体验，被认为是随着时间的推移和通过锻炼的体验而习得的联想（Conroy & Berry，2017）。然而，一些研究者似乎更喜欢使用内隐态度（Calitri, Lowe, Eves, & Bennett, 2009; Markland, Hall, Duncan, & Simatovic, 2015）或情感联想（Sala, Baldwin, & Williams, 2016）来表述自动情感评价，但这些术语的同义用法通常是不被接受的。在情感和健康行为体系的背景下，自动情感评价最类似于 IBC 模型中内隐态度的自动情感处理方式，它反映了对身体活动的瞬时情感评价，并可以从对身体活动的直接或替代体验中获得（比如在身体活动中观察他人时所唤醒的情感；Williams & Evans, 2014）。由 Conroy 和 Berry（2017）提出的强调自动情感评价作用的身体活动双系统模型，认为愉快的情感过程提供了内部奖励以激励身体活动（近端效应）。另外，自动情感评价不同于熟虑情感评价，但两种评价可能存在重叠。自动情感评价可以快速、不费力的进行，不需要有意识的处理或精细化，既可以影响自动动机（比如无意识的促进目标追求），也可以影响熟虑情感过程（比如情感预期或情感态度；Kiviniemi, Voss-Humke, & Seifert, 2007），而这些过程被认为是决定锻炼动机的上游因素。

图 1 总结了如何在双系统体系内将情境线索、自动情感评价与身体活动联系起来的概念模型。具体而言，自动情感评价是指人们的经验随时间的推移不断积累的过程中，个人根据身体活动的概念和不同的情感体验之间的特殊联系来创造记忆，这些记忆的特征组成了联想的效价和强度。当一个人接触到激活身体活动概念的线索时，相关的记忆被激活，这种激活方式可以扩散到网络中密切相关的节点（比如特征情感体验）。实际上，在没有必要经历明显的快乐或不快乐的情况下，可能会产生一种直觉使得平衡偏向或偏离身体活动。因此，对身体活动的自动情感评价可能会产生一种冲动，从而增加不活跃的个体参与锻炼的可能性。另外，联想-命题评价（Associative-Propositional Evaluation, APE）模型主要分析了内隐态度转变为外显态度的心理机制，它认为大脑存在两种信息处理方式，即联想和推理评价过程（Gawronski & Bodenhausen, 2006）。内隐态度是联想评价过程，由记忆中的联结唤起；外显态度是推理评价过程，以对真值（即价值和信念）的确认作为特征（叶娜 & 佐斌，2007）。APE 模型在这里被用来解释自动情感评价与 TPB 的态度构念之间的联系。

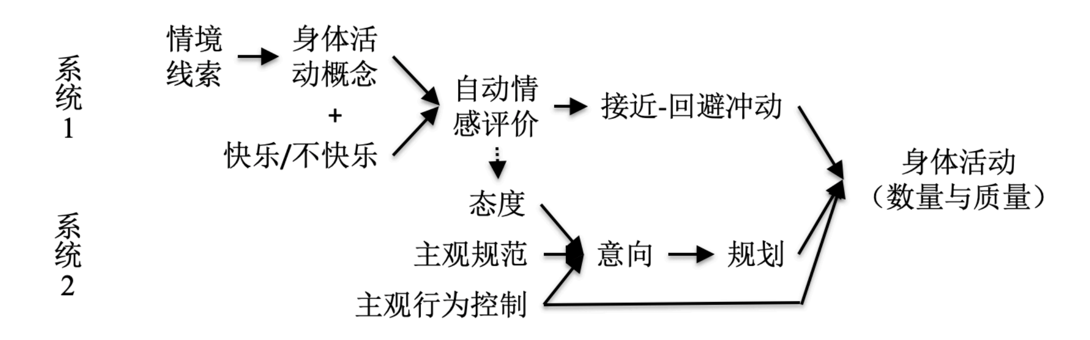


图 1　突出自动情感评价作用的身体活动双系统概念模型

资料来源：Conroy 和 Berry (2017)

此外，经常锻炼者比不经常锻炼者拥有更积极的自动情感评价，而积极的自动情感评价具有对未来锻炼行为的预测能力。Antoniewicz 和 Brand（2016b）的研究表明，改变自动情感评价后，锻炼行为会出现即刻差异，而另有研究采用了 1 周（Calitri et al., 2009）、3 个月（Antoniewicz & Brand, 2016a）或 6 个月的时间范畴（Endrighi et al., 2016），同样发

现自动情感评价与短期和长期锻炼行为都有着联系。并且，自动情感评价越积极的锻炼者越有可能达到理想的锻炼频率（Brand & Antoniewicz, 2016）。另外，现有研究为将自动情感评价作为目标的行为改变干预的可行性提供了初步建议（Antoniewicz & Brand, 2014; Calitri et al., 2009; Endrighi et al., 2016）。但在将自动情感评价作为目标进行锻炼干预之前，首先要明确自动情感评价与锻炼行为之间是否具有因果关系。而目前仅有一项研究证实了两者之间确实存在着因果关系（Antoniewicz & Brand, 2016b）。还有研究报道了自动情感评价和自我报告的锻炼量之间存在正相关（Bluemke, Brand, Schweizer, & Kahlert, 2010; Chevance, Caudroit, Romain, & Boiche, 2017）。

2.2 身体活动双系统模型中系统 2 的构建

早期的锻炼行为理论是基于信息加工范式的认知主义理论，它们出自于同一元理论体系，关注的是意向和规划这样的认知构念，而忽略了情感构念的动机属性以及自动化过程的重要性（Ekkekakis, 2017）。Biddle 等人（2007）建议将锻炼行为理论分为五类：第一，信念-态度理论关注的是行为意向的认知前提，即某人准备为实现目标行为而投入的努力，比如 TPB。第二，能力基础理论主要以自我效能构念为例，自我效能被定义为个人对他人组织和执行达到指定类型的表现所需的行动过程的能力的判断。第三，控制基础理论认为个人有体验自己作为行为的发起者和调节者的内在欲望或目标，比如自我决定理论将这种渴望归因于对自我决定的基本心理需求。第四，以跨理论模型为代表的阶段模型则将行为改变概念化为使一个人更接近所设想的目标的过程。第五，以健康行动过程取向理论为代表的混合模型将阶段概念与动机变量相结合，从而预测意向，并添加了决策后变量（比如执行意向）。这些理论模型定义了研究中的变量，提供了变量之间的结构，描述了应该如何操作变量的假设，允许研究的复制和泛化，为假设的检验和证伪提供了对话框架。随后在研究中发现，这些理论都有各自的优点与缺点，没有哪一个理论能够全面且便于操作地对锻炼行为做出解释、预测，并能有效地采取干预措施（沈梦英，毛志雄，张一民，2010）。鉴于这些理论本身存在的一些不足和待改进之处，不少国内外研究者尝试将多种理论结合起来构建一个整合模型，以期提高模型对锻炼行为的解释、预测能力（冯玉娟，毛志雄，2014；Hagger & Chatzisarantis, 2014）。

2.3 身体活动双系统模型的控制原则

在模型的控制方面，基于 Strack 和 Deutsch（2004）建议的熟虑-冲动模型（Reflective-Impulsive Model, RIM），Bluemke 等人（2010）以跑步为例，提出了基于竞争控制的身体活动 RIM。它认为如果熟虑和冲动系统的方向是一致的，那么两个系统就会增强跑步的执行；如果两个系统的方向是冲突的，那么它们就会竞争控制跑步行为。而 Aarts 等人（1997）早前提出了基于协调控制的身体活动与习惯养成模型，将主要涉及到的对过往锻炼体验记忆的认知过程整合到一个锻炼习惯养成模型中。它假设锻炼的起始在很大程度上是由熟虑决策决定的，随着不断的重复和更多的练习，最终形成了一种习惯。此外，Rhodes 和 Rebar（2018）新近提出基于层级控制的双系统模型，它认为像身体活动这样的复杂行为是分层描述每个行动的，这些行动是由较低层级的子行动（sub-actions）所组成的，行动之间存在熟虑转换，而其子行动之间又存在自动转换。

2.3.1 竞争控制

RIM 强调了与社会行为的联系，将熟虑和冲动看作是行为的过程（Strack & Deutsch, 2004）。它关注的是由线索或刺激激活的评价和语义联系。这些联系是在接近或逃避导向的冲动系统中对刺激情感反应的基础。熟虑系统将这些自动联想阐述为命题，联想的真值是经过熟虑后确定的。它可以克服习惯性反应，或在习惯失效的新情况下，将行动规划整合到一起。RIM 假定信息在两个分离的系统中并行加工，经过多种渠道相互联系。熟虑系统的作用路径：行为选择→形成行为意向→启动行为计划→行为意向有助于执行行为；冲动系统的作用路径：知觉到刺激→激发冲动系统网络中的更多节点→激活行为模式（许昭，

毛志雄，2015）。依据 RIM，没有一个过程是完全熟虑或冲动的，两个系统相互作用影响行为。尽管 RIM 与 APE 模型分享某些重要的假设，两个模型之间也有明显的区别。例如，RIM 假设这两个系统是同时被独立调用的。而在 APE 模型的默认干预逻辑中，内隐联想为进一步的命题加工提供了默认值，使所有的系统 2 加工都依赖于初始的系统 1 输入（Evans & Stanovich, 2013）。

Bluemke 等人（2010）以跑步为例，认为如果熟虑系统和冲动系统的方向是一致的，人们充分了解跑步的益处，并且跑步与正效价情感形成情感联系，那么两个系统就会增强跑步的执行；如果两个系统的方向是冲突的，人们充分了解跑步的益处，但将跑步知觉为不快乐的，把跑步与负效价情感形成情感联系，那么两个系统就会竞争控制跑步行为。此外，许昭和毛志雄（2015）构建了由 10 个变量的熟虑系统和 1 个变量的冲动系统组成的身体活动 RIM 模型，经模型验证发现，双系统对身体活动的影响受行为习惯性（有无规律运动训练经历）的调节，行为习惯性越强，冲动作用越小，反之越大；另外，双系统对身体活动的影响还受意志品质（执行熟虑系统的能力）的调节，意志品质越高，熟虑系统对身体活动的影响越大，反之冲动系统的作用更大，认为身体活动是由两个系统共同决定的。

2.3.2 协调控制

Aarts 等人（1997）提出的身体活动与习惯养成模型，将主要涉及过往锻炼体验记忆的认知过程（对锻炼习惯决定因素的认识）整合到一个锻炼习惯养成模型中（图 2）。并且认为该模型只是包括社会-人口、生物和遗传因素的一般模式中的一部分，强调了影响身体活动的心理因素（近端效应）。身体活动与习惯养成模型假设锻炼的起始在很大程度上是由熟虑决策决定的，比如态度、主观社会规范、行为控制和运动经验（图 2 上半部分描述了起始的熟虑决策过程）；随着不断的重复和更多的练习，锻炼的基本决策过程逐步从启发式的决策过程转变为自动化的决策过程，最终形成了一种不再需要理性思考指导的习惯（图 2 左下角的粗体箭头表示习惯养成过程的反馈循环）。

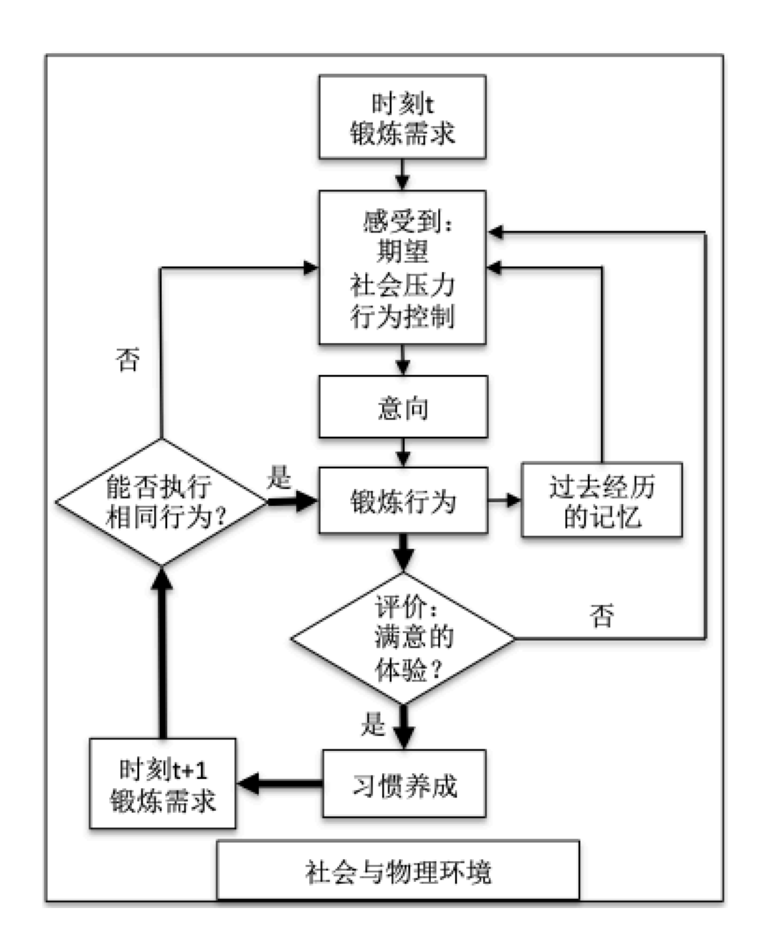


图 2　身体活动与习惯养成模型
资料来源：Aarts et al. (1997)

然而理性行为和习惯可以看作一个统一体的两端。在目标定向行为既不完全是自动化的，也不完全是熟虑的情况下，决策者可以使用类似于启发式决策的方法来进行选择。可以将这种启发式的决策过程设想为一种认知捷径。例如，不仔细检查所有执行行为的后果，而是迅速考虑行为的最突出优势，以明确它的可行性，或简单地确认能否执行相同的行为。随着练习的增加，熟虑的决策过程可能会通过对过去行为经历的记忆逐步转变为更具启发性的过程（认知捷径的发展）。一旦这些结果储存在记忆中，并能够很容易从记忆中被检索，就不再需要反复思考。启发式处理的核心特征就是从记忆中检索过去的决策过程中得出的结论或作出的决定。在身体活动与习惯养成模型中，这种学习表现为过去行为对记忆的反馈以及运动事件的记忆对感知的影响。

2.3.3　层级控制

身体活动并不是一种简单的行为，而是由许多子行动组成的各种复杂行为的组合（Rhodes & Rebar, 2018）。虽然习惯与熟虑的动机可能是一种全或无的现象（即一种影响只能解释某一时刻的行为），但这不一定适用于长时间（30 分钟以上）的身体活动。Gardner 等人（2016）基于 Cooper 和 Shallice（2006）的理论，从行动-阶段的角度概述了这一过程，认为身体活动这样的复杂行为是分层描述每个行动的，这些行动由较低层级的子行动组成。这种理解身体活动的方式允许各种行动序列模块化（chunked）为自动调节的行为（Graybiel, 2008）。如图 3 所示，一个刚开始跑步锻炼的新手需要预先仔细考虑身体活动的各个方面，从准备决策（时间、服装等的选择）到执行方面（路线、跑步速度、节奏和风格）。随着时

间的推移，通过简单的子行动（跑步风格）的技能获取，或更高层级决策与行动的习惯养成，这些方面中的某一部分可能会形成自动化（前往目的地，决定要做的活动）。随后当锻炼者形成将前一个子行动的结束与下一个子行动的开始相联系的记忆时，每个子行动将不再需要思考，而是由情境触发的接近倾向自动提示行动。

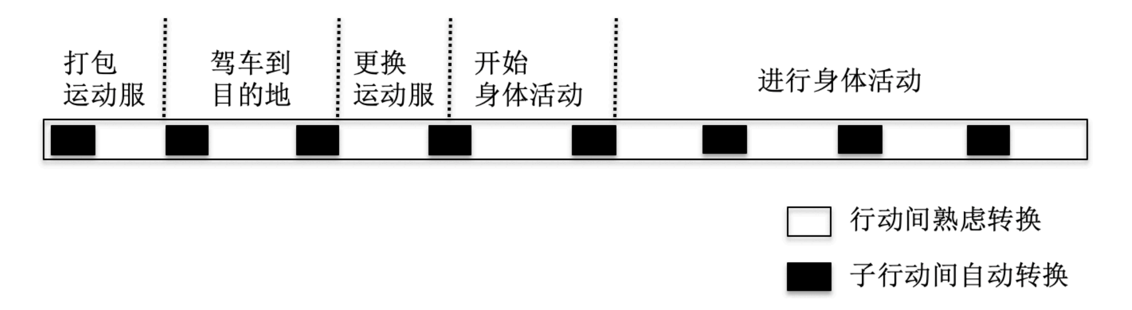


图 3　熟虑与习惯促进身体活动之间的建议转换
资料来源：Rhodes 和 Rebar (2018)

采用这种方法来理解锻炼习惯的形成需要识别行为序列的要素。考虑到锻炼决策（做出锻炼的决定）是影响锻炼行为的重要因素，Verplanken 和 Melkevik（2008）提出了一种更为可行的方法，即在身体活动的启动阶段测量习惯。随后 Gardner 和同事（Gardner, Phillips, & Judah, 2016; Phillips & Gardner, 2016）提出启动/选择阶段（决定对其他潜在刺激采取行动）与执行阶段（后续的有序行动）可能是对复杂身体活动序列概念化的有效途径。Kaushal 等人（2017）也认为准备阶段（身体活动前的行为和启动）与执行阶段可能是理解个人锻炼习惯的有效途径。虽然启动阶段与执行阶段都可能成为习惯，但启动阶段对于理解有规律的身体活动可能更重要，因为它代表了之前的决策过程。相比之下，执行阶段可以解释身体活动的持续时间或努力程度，但似乎不能解释身体活动被反复选择和启动的原因（Gardner et al., 2016）。因此，上述研究表明启动阶段可能才是身体活动参与频率的主导预测因子。

总之，在模型的控制方面，上述三个模型所建议的竞争、协调与层级控制的原则，为系统 1 与系统 2 在复杂身体活动中执行有序控制提供了合理解释。另外，协调控制是一种序列加工的方式，而竞争控制是双系统并行加工的方式。鉴于竞争控制并不符合人类认知加工的经济、快捷原则，在心理学上不太可能实现（艾炎，胡竹菁，2018）。因此，相较于 RIM 提出的竞争控制，身体活动更可能是一种相互协作、分层执行的复杂行动组合。

3　强化学习视角下的双系统模型

近年来机器学习领域的计算 RL 理论对心理学和神经科学产生了持久而深远的影响（Botvinick, Niv, & Barto, 2009; Botvinick & Weinstein, 2014）。虽然 RL 最初被视为特定计算技术的存储库，但它已逐步发展成为用于思考人与动物的动机行为和学习过程的一个通用框架。这些影响最初体现在对经典条件反射和操作条件反射的研究中，随后扩展到运用时序差分学习范式解释中脑多巴胺神经元活动的一个框架。目前基于 RL 的理论已被用于解决各类问题。因此，有必要从 RL 的角度分析双系统模型的构建与控制问题。在模型的构建方面，基于认知神经科学的观点，无模型（model-free, MF）和基于模型（model-based, MB）的 RL 分别表示双系统的系统 1 与系统 2。在模型的控制方面，Dyna 协作架构假设 MF 受到 MB 控制，建议通过目标控制不断影响习惯过程，促进与目标一致的习惯养成。另外，分层 RL（hierarchical RL, HRL）假设 MF 与 MB 控制的组合方式应该是任意的，这进一步

扩展了身体活动双系统模型的层级控制机制。并且，MF 的目标选择为锻炼习惯养成的重要性提供了依据，有强烈锻炼习惯的人会优先考虑他们熟悉的行为方式。

3.1 强化学习概述

RL 是指智能体（agent）在与环境的连续互动过程中学习最优行动策略的机器学习问题以及解决这类问题的方法（Sutton & Barto, 2018）。在 RL 任务中有两个互动的对象：智能体与环境。一般而言，学习者或决策者被称为智能体，它感知外界环境的状态和反馈的奖励，并进行学习与决策。智能体泛指人、动物以及机器人和自动驾驶汽车等人工智能应用。环境是指智能体外部的所有事物，它受到智能体行为的影响而发生状态变化，并反馈给智能体对应的奖励。如图 4 所示，在每一个时间步 t，智能体从环境中观察到一个状态 S_t 与一个奖励 R_t，采取一个行动 A_t。环境根据智能体的决策行动决定下一步 t+1 的状态 S_{t+1} 与奖励 R_{t+1}。学习的策略表示成给定状态下采取的行动。智能体的目标不是短期奖励的最大化，而是长期累积奖励的最大化。

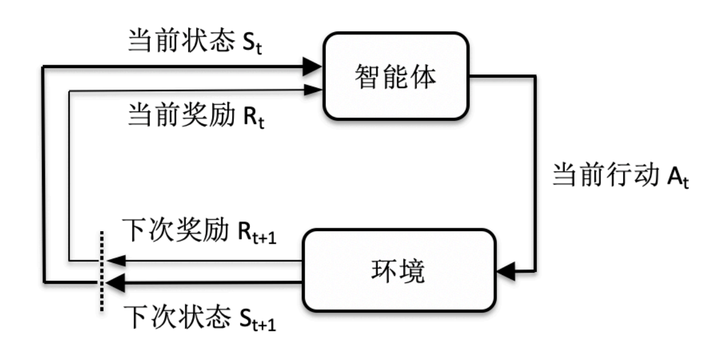


图 4 强化学习中智能体-环境的互动
资料来源：Sutton 和 Barto (2018)

3.2 强化学习视角下双系统模型的构建

RL 的核心问题是对状态-行动配对（预期贴现回报）的值估计，用以指导行动的选择。一般而言，解决这个问题有两种策略：一种是根据经验进行值估计的 MF 策略，另一种是根据环境模型（奖励和转换函数）学习的 MB 策略，随后用它来规划最优策略。MF 策略由习惯系统执行，MB 策略由目标定向系统执行（Daw, Markman, & Otto, 2005; Dolan & Dayan, 2013）。

从心理学的角度看，MF-RL 是桑代克的效应定律的衍生物，个体依据这个定律习惯性重复过去强化过的行为（Gershman, Joel, & Dayan, 2014）。它不依赖于环境的内部模型，相反，配对值以缓存格式（比如查找表）存储，允许对其快速检索。这些值可以使用简单的试错驱动的学习规则（比如时序差分学习范式）进行增量更新（Sutton & Barto, 2018）。MF 策略的主要缺点是缺乏灵活性：当环境或任务发生改变时，需要重新学习整个缓存值的集。由于这种不灵活重复导致 MF 策略习惯化。因此，MF-RL 从理论上捕获了习惯的关键属性——对奖励变化的不敏感。

相比之下，MB-RL 是托尔曼的认知地图概念的衍生物。MB 策略以内部模型的形式表示其认知，当发生改变时，可以在局部对其进行修改（只需修改模型的一部分）。因此，与 MF 策略不同，MB 策略不需要缓存值。它可以在不重新学习整个模型的情况下，灵活地修改策略以追求目标。然而执行 MB 策略不可避免地比 MF 策略需要更多的时间和资源（Daw et al., 2005; Keramati, Dezfouli, & Piray, 2011）。另外，TPB 将客观环境的作用（比如来自配偶、亲友、榜样的社会支持与锻炼的物质条件）体现在主观规范与主观行为控制两个因素

中。它被认为联通了个体与环境，将客观社会环境与物质条件等因素对个体的约束转变为个体的主观感知，解释了客观环境因素对行为的作用机制（沈梦英 等，2010）。这类似于 MB-RL 中内部模型的作用。

3.3 强化学习视角下双系统模型的控制机制

当学习一个新的不确定任务时，个体应该仔细考虑行动的后果（MB），但是在对一个稳定的任务进行长时间的练习之后，个体通常可以通过重复同一个任务来更快地获得相同的结果（MF）。尽管在最初的双系统 RL 中假设了一种竞争的赢者通吃机制，认为 MF 或 MB 系统获得对行为的控制是取决于哪个系统提供了更可靠的行动值估计（Daw et al., 2005; Daw, Gershman, Seymour, Dayan, & Dolan, 2011）。然而 RL 模型调用的这种并行处理作为一种昂贵的中央处理，几乎随时准备执行规划与熟虑，因而在心理学上不太可能实现（Evans & Stanovich, 2013）。后续研究更倾向于一种动态整合，两个系统都根据它们的相对可靠性来计算行动值（Lee, Shimojo, & O'Doherty, 2014）。因此，结合 MF 和 MB 的模型不是简单地假设行为是习惯或目标定向的，而是在不同的行为控制模型的互动中探索如何产生适应行为。

3.3.1 Dyna 协作架构

除整合独立计算的行动值以外，一些 RL 模型还建议了 MB 规划对 MF 学习的直接影响。例如，Gershman 等人（2014）发现在两步决策任务中，受试者在独立学习了第二步奖励后，在第一步任务中改变了他们的选择偏好（被认为反映了 MF 习惯处理）。为了解释这一发现，Gershman 等人提出 MB 系统模拟了一个完整的两步决策过程，MF 系统可以从模拟中学习。并且，他们还引入了一个能够有效整合 MF 与 MB 系统的 Dyna 架构。如图 5 所示，Dyna 架构中行动完全由 MF 系统控制（允许以最小的计算成本在线进行决策），MB 系统通过离线的方式训练 MF 系统（在两段真实体验之间），可以对行动产生间接影响。因此，Dyna 架构中使用模拟体验替代了部分真实体验。具体而言，MB 系统回放经历的状态-行动配对，然后在习得的环境模型基础上模拟一次转换和奖励。随后 MF 系统从这些模拟轨迹中进行学习，如同从真实体验中学习一样。

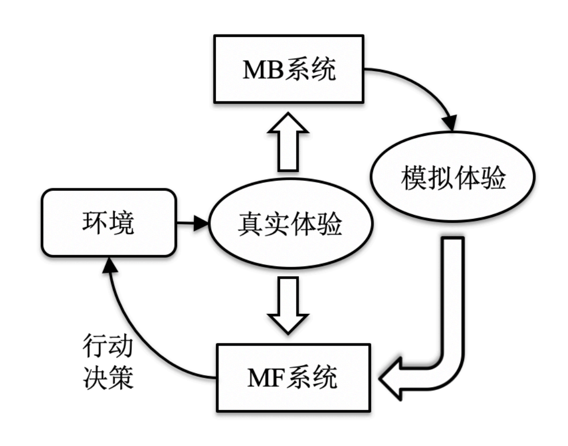


图 5 Dyna 协作架构
资料来源：Gershman et al. (2014)

Dyna 架构假设熟虑的学习规划能够更新与改进习惯学习的值估计，通过目标影响习惯学习表示的行动值，促进与目标一致的习惯养成。因此，不同于我们通常认为的习惯是 MF 学习结果的 RL 假设，习惯是习得的行为模式，受到 MB 控制（Dezfouli & Balleine, 2012）。这可能是由于标准的两步 RL 决策任务无法有效捕获对结果不敏感的习惯过程，而在这些

任务中的 MF 学习可能会反映其他刺激驱动的策略，比如简单的启发式决策。今后对 MF 学习的研究可能需要开发新的实验任务，以便更有效地捕获习惯过程。

3.3.2 分层强化学习

HRL 的主要目标是将复杂的问题分解成多个小问题，通过分别解决小问题从而达到解决原问题的目的（周文吉，俞扬，2017）。HRL 的核心思想是引入抽象机制对学习任务进行分解。Sutton 等人（1999）的研究中，学习任务被抽象为若干选择项（options），并将这些选择项作为一个特定的行动加入到原来的行动集中。一个选择项可以理解为要完成某子目标而按一定策略执行的行动序列。选择项可以由设计者根据专家知识预先确定，利用先验知识加速从学习到相关任务的迁移，也可以通过 MF 方法学习（Botvinick et al., 2009），或重复连接到一个模块化的行动序列中（Dezfouli & Balleine, 2012）。因此，HRL 可以划分为 MF-HRL 和 MB-HRL。

HRL 理论假设将单个动作串联起来，并将序列作为一个反应单元或模块可能是有利的（Botvinick & Weinstein, 2014）。因为通过模型化方式执行的快速反应会带来更多的平均回报（Keramati et al., 2011）。任何层级的决策原则上都可以通过 MF 或 MB（或两者任意组合）的方式执行。例如，可以通过目标控制在较高层级上选择先前习得的行动序列来执行期望的目标（Dezfouli & Balleine, 2013），也可以执行对 MB 规划的 MF 控制。Cushman 和 Morris（2015）研究表明，个人通过习惯控制来解决目标选择过程，同时利用规划实现被选择的目标。如图 6 所示，左图描述了一个有大量可能的未来行动序列树的程式化选择问题，它为 MB 规划提供了一个困难的搜索问题。右图描述的是 Cushman 和 Morris（2015）的假设，MB 规划可以通过初始的 MF 进行目标选择来简化步骤（三个连续实线箭头）。这些模型不是互相排斥的，而是基于一个共同的假设：个人会灵活地运用习惯或目标控制，跨层级组合行为以适应特定任务的需求。

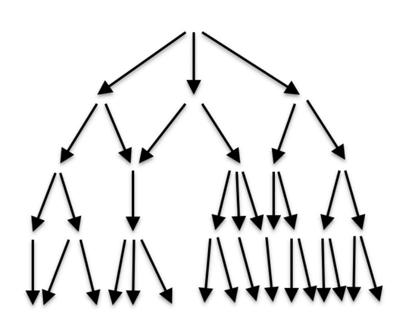
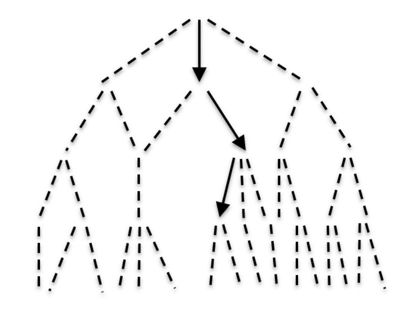



图 6 基于模型的搜索与无模型的目标选择
资料来源：Daw (2015)

这种 MF 的目标选择也符合个人一贯的行为方式。出于习惯，头脑中会减少已有的刻意反应，甚至在需要作出明确决策时也会减少关注。例如，一个多属性选择任务涉及一系列出行方式的决策，有较强骑自行车或开车习惯的人进行了有限的信息检索，只考虑较少的出行方案，并偏向于他们习惯的选择（Aarts, Verplanken, & Van Knippenberg, 1997; Betsch et al., 2001; Verplanken, Aarts, & Van Knippenberg, 1997）。从本质上讲，有强烈习惯的人处理信息的方式减少了他们考虑采取其他行动的可能性。

3.4 强化学习视角下锻炼者-体育环境的互动模式

从 RL 的角度看，锻炼者通过锻炼行为作用于体育环境，而体育环境又依据锻炼者的行为变化进行状态更新，并反馈给锻炼者相应的奖励，促使锻炼者不断调整锻炼行为，以期累积获得最高回报（图 7）。首先，锻炼者可以通过双系统控制自身的身体活动：MF 系

统（系统 1）受到情境线索提示，激活包括自动情感评价在内的锻炼习惯行为；MB 系统（系统 2）通过主观规范与主观行为控制将客观体育环境的作用转变为锻炼者的主观感知（即内部模型，未画出）。系统 2 的加工依赖于系统 1 的输入，比如态度的情感成分受到自动情感评价的影响（见图 7 由系统 1 指向系统 2 的虚线箭头）；通过系统 2 对系统 1 的作用，促进与目标一致的习惯养成（见图 7 由系统 2 指向系统 1 的虚线箭头）。此外，双系统是以一种相互协作的形式对身体活动执行分层控制的。例如，在开始一项新的身体活动或在不同子行动之间进行转换时，锻炼者更多依赖于系统 2，但在特定情境下重复已习得的锻炼习惯行为时启用系统 1，从而达到节省资源、提高效率的目的。

其次，特定的体育环境能够自动激活锻炼习惯行为，被认为是锻炼习惯自动化的情境线索。最近有学者将体育环境划分为建成环境、社会环境和信息通讯环境（张展嘉 等，2018）。相关研究表明，从城市的步行性、锻炼者的社交网络以及智能手机使用的信息反馈这三个方面来看，体育环境能够显著影响锻炼行为（Althoff et al., 2017; Aral & Nicolaides, 2017）。另外，对锻炼行为的奖励通常可以划分为内部和外部奖励。区别于外部奖励（报酬），内部奖励被认为是情绪反应和随后对身体活动的情感判断，代表了个人对情感体验的期望，有助于形成更强烈的习惯，比如对内在调节、享受、乐趣和愉悦感的期望（Ekkekakis, Hargreaves, & Parfitt, 2013; Rhodes, Fiala, & Conner, 2009）。外部奖励又与主观规范有关，可以通过现代媒介化社会中起主导作用的互联网和智能手机等新媒介的积极运用来实现（张业安，2018）。

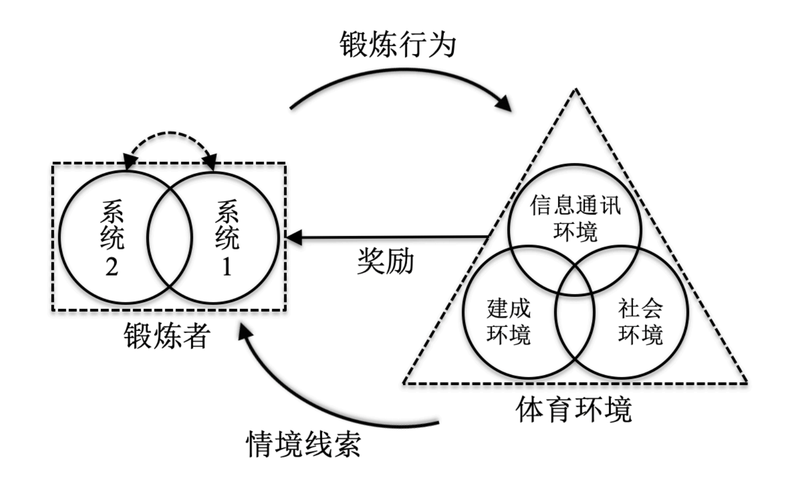


图 7　强化学习视角下锻炼者-体育环境的互动

此外，如何在 RL 框架内运用奖励机制来激发个人的锻炼行为，或许是行为改变干预策略设计的一个潜在目标。MF-RL 是依据候选行动的奖励历史为其分配存储值的，由此形成类似于刺激-应答习惯的存储策略。可以认为 MF-RL 维护了一个查找表，包含对每个状态-行动配对的（存储值）未来回报的预测。另一方面，内隐评价依赖于 MF 学习过程（Kurdi, Gershman, & Banaji, 2019）。如果先前的锻炼体验让人感觉良好，那么锻炼的概念和记忆中的情感状态之间就会产生联想。随后，MB-RL 将这些自动联想阐述成命题（其中联想的真值是经过熟虑后确定的），这也符合 APE 模型的假设（Conroy & Berry, 2017）。这进一步提示我们，如果锻炼者对不同“情境线索-锻炼习惯配对”的相关记忆进行主观赋值，那么同样可以构建一个内部的锻炼习惯查找表，为从 RL 角度理解个人锻炼决策的动机效应提供了新的见解。例如，当个人拥有对锻炼的积极熟虑评价（比如对健康的认识、个人能力的评价、可及社会支持的感知）和愉快的情感体验（比如一种生理兴奋、对过去成绩的自豪感），因而对身体活动的赋值较高，那么会更倾向于选择积极的身体活动；相反，如果个人

对身体活动持有负面评价，背负着不愉快的体验（比如身体疼痛、对运动损伤的恐惧、由羞愧带来的情感伤害），那么会更倾向于久坐不动的生活方式，因为久坐不动的行为也会伴有愉快的体验（比如使用社交软件、上网冲浪、看电视），导致对久坐行为的赋值高于持有负面评价的身体活动。

总之，RL 视角下锻炼者与体育环境的互动模式强调了两者之间的连续互动，这与社会生态系统理论既有联系又有区别。例如，由 Bronfenbrenner（1979）提出的生态系统理论，作为社会生态模型早期的版本，强调了个体的发展嵌套于相互影响的一系列环境系统之中，在这个系统中环境与个体互相作用并最终影响个体的发展。研究者最近通过系统评述身体活动中社会认知与建成环境的互动机制，发现两者只在休闲体育活动中存在互动（Rhodes, Saelens, & Sauvage-Mar, 2018）。这或许提示我们，应当更多关注休闲体育活动领域中体育环境对锻炼行为的影响。

4　小结

身体活动的双系统理论认为情境线索是通过快速、自动、无意识的认知过程（习惯）影响锻炼行为的，这不仅弥补了早期锻炼行为理论模型的不足，也在体育环境与个人锻炼习惯的养成之间构建了一条合理的路径。并且，在双系统模型中通过自动情感评价将情境线索与身体活动联系起来，可以更好地解释情感启发下的复杂锻炼行为。此外，身体活动可能是一种相互协作、分层执行的复杂行动组合。习惯与熟虑系统交替控制身体活动，从而使锻炼者与体育环境之间实现连续的互动。RL 视角下的 Dyna 协作架构建议习惯养成依赖于目标控制，而情境线索也通过自动化过程影响身体活动，因此认为身体活动的自动化过程（习惯）可能是整合情境线索（环境）与目标定向控制过程（意向）的关键：自动化过程为目标定向控制提供情境线索输入，同时受到目标定向控制的影响。最后提出 RL 视角下锻炼者-体育环境的互动模式，以期为今后锻炼行为的研究提供一个全新的视角。

参考文献

艾炎，胡竹菁. (2018). 推理判断中双重加工过程的协作与转换机制. *心理科学进展, 26*(10), 1794–1806.

段文婷，江光荣. (2008). 计划行为理论述评. *心理科学进展, 16*(2), 315–320.

冯玉娟，毛志雄. (2014). 高中生身体活动意向和行为的促进策略：自我决定动机对 TPB 的贡献. *体育科学, 34*(8), 64–69.

郭强，汪晓赞，蒋健保. (2017). 我国儿童青少年身体活动与久坐行为模式特征的研究. *体育科学, 37*(7), 17–21.

韩慧，郑家鲲. (2016). 西方国家青少年体力活动相关研究述评——基于社会生态学视角的分析. *体育科学, 36*(5), 62–70.

沈梦英，毛志雄，张一民. (2010). 中国成年人锻炼行为的影响因素——HAPA 与 TPB 两个理论模型的整合. *体育科学, 30*(12), 48–54.

司琦. (2007). 身体活动的行为科学理论综述. *体育科学, 27*(9), 72–80.

许昭，毛志雄. (2015). 身体活动熟虑-冲动双系统模型的构建与检验. *体育科学, 35*(8), 16–23.

叶娜，佐斌. (2007). 联想-命题评价模型——态度改变的新解释. *心理科学进展, 15*(5), 834–839.

张加林，唐炎，胡月英. (2017). 我国儿童青少年体育环境特征与存在问题研究. *体育科学, 37*(3), 21–34.

张业安. (2018). 青少年体质健康促进的媒介责任：概念、目标及机制. *体育科学, 38*(6), 14–26.

张展嘉，王正珍，于洪军，陈蔚云. (2018). 第 65 届美国运动医学会年会关于身体活动促进的研

究热点与进展综述. *北京体育大学学报, 41*(8), 72–76.
周文吉，俞扬. (2017). 分层强化学习综述. *智能系统学报, 12*(5), 590–594.
朱为模. (2009). 从进化论、社会-生态学角度谈环境、步行与健康. *体育科研, 30*(5), 12–16.
Aarts, H., Paulussen, T., & Schaalma, H. (1997). Physical exercise habit: On the conceptualization and formation of habitual health behaviours. *Health Education Research, 12*, 363–374.
Aarts, H., Verplanken, B., & Van Knippenberg, A. (1997). Habit and information use in travel mode choices. *Acta Psychologica, 96*, 1–14.
Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior & Human Decision Processes, 50*, 179–211.
Althoff, T., Sosic, R., Hicks, J. L., King, A. C., Delp, S. L., & Leskovec, J. (2017). Large-scale physical activity data reveal worldwide activity inequality. *Nature, 547*, 336–339.
Antoniewicz, F., & Brand, R. (2014). Automatic evaluations and exercise setting preference in frequent exercisers. *Journal of Sport & Exercise Psychology, 36*, 631–636.
Antoniewicz, F., & Brand, R. (2016a). Learning to like exercising: Evaluative conditioning changes automatic evaluations of exercising and influences subsequent exercising behavior. *Journal of Sport & Exercise Psychology, 38*, 138–148.
Antoniewicz, F., & Brand, R. (2016b). Dropping out or keeping up? Early-dropouts, late-dropouts, and maintainers differ in their automatic evaluations of exercise already before a 14-week exercise course. *Frontiers in Psychology, 7*, 838.
Aral, S., & Nicolaides, C. (2017). Exercise contagion in a global social network. *Nature Communication, 8*, 14753.
Betsch, T., Haberstroh, S., Glockner, A., Haar, T., & Fiedler, K. (2001). The effects of routine strength on adaptation and information search in recurrent decision making. *Organizational Behavior and Human Decision Processes, 84*, 23–53.
Biddle, S. J. H., Hagger, M. S., Chatzisarantis, N. L. D., & Lippke, S. (2007). Theoretical frameworks in exercise psychology. In G. Tenenbaum & R. Eklund (Eds.). *Handbook of Sport Psychology* (3rd ed., pp. 537–559). Hoboken, NJ: Wiley.
Bluemke, M., Brand, R., Schweizer, G., & Kahlert, D. (2010). Exercise might be good for me, but I don’t feel good about it: Do automatic associations predict exercise behavior? *Journal of Sport & Exercise Psychology, 32*, 137–153.
Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition, 113*, 262–280.
Botvinick, M. M., & Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. *Philosophical Transactions of the Royal Society B, 369*, 20130480.
Brand, R., & Antoniewicz, F. (2016). Affective evaluations of exercising: The role of automatic-reflective evaluation discrepancy. *Journal of Sport & Exercise Psychology, 38*, 631–638.
Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, Massachusetts: Harvard University Press.
Calitri, R., Lowe, R., Eves, F. F., & Bennett, P. (2009). Associations between visual attention, implicit and explicit attitude and behaviour for physical activity. *Psychology & Health, 24*, 1105–1123.
Chevance, G., Caudroit, J., Romain, A. J., & Boiche, J. (2017). The adoption of physical activity and

eating behaviors among persons with obesity and in the general population: The role of implicit attitudes within the Theory of Planned Behavior. *Psychology, Health & Medicine, 22*, 319–324.

Conroy, D. E., & Berry, T. R. (2017). Automatic affective evaluations of physical activity. *Exercise & Sport Sciences Reviews, 45*, 230–237.

Cooper, R. P., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behaviour. *Psychological Review, 113*, 887–916.

Cushman, F., & Morris, A. (2015). Habitual control of goal selection in humans. *Proceedings of the National Academy of Science of the United States of America, 112*, 13817–13822.

Daw, N. D., Markman, A. B., & Otto, A. R. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*, 1704–1711.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*, 1204–1215.

Daw, N. D. (2015). Of goals and habits. *Proceedings of the National Academy of Science of the United States of America, 112*, 13749–13750.

Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience, 35*, 1036–1051.

Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Computational Biology, 9*, e1003364.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron, 80*, 312–325.

Ekkekakis, P. (2017). People have feelings! Exercise psychology in paradigmatic transition. *Current Opinion in Psychology, 16*, 84–88.

Ekkekakis, P., Hargreaves, E. A., & Parfitt, G. (2013). Envisioning the next fifty years of research on the exercise-affect relationship. *Psychology of Sport & Exercise, 14*, 751–758.

Endrighi, R., Basen-Engquist, K., Szeto, E., Perkins, H., Baum, G., Cox-Martin, M., … Waters, A. J. (2016). Self-reported and automatic cognitions are associated with exercise behavior in cancer survivors. *Health Psychology, 35*, 824–828.

Evans, J., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science, 8*, 223–241.

Gardner, B., De Bruijn, G.-J., & Lally, P. (2011). A systematic review and meta-analysis of applications of the self-report habit index to nutrition and physical activity behaviours. *Annals of Behavioral Medicine, 42*, 174–187.

Gardner, B., Phillips, L. A., & Judah, G. (2016). Habitual instigation and habitual execution: Definition, measurement, and effects on behaviour frequency. *British Journal of Health Psychology, 21*, 613–630.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731.

Gershman, S. J., Joel, D., & Dayan, P. (2014). Retrospective revaluation in sequential decision making: A tale of two systems. *Journal of Experimental Psychology: General, 143*, 182–194.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience, 31*,

359–387.
Hagger, M. S., & Chatzisarantis, N. L. D. (2014). An integrated behavior change model for physical activity. *Exercise & Sport Sciences Reviews, 42*, 62–69.
Kaushal, N., Rhodes, R. E., Meldrum, J., & Spence, J. C. (2017). The role of habit in different phases of exercise. *British Journal of Health Psychology, 22*, 429–448.
Keatley, D. A., Clarke, D. D., & Hagger, M. S. (2012). Investigating the predictive validity of implicit and explicit measures of motivation on condom use, physical activity, and healthy eating. *Psychology & Health, 27*, 550–569.
Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology, 7*, e1002055.
Kiviniemi, M. T., Voss-Humke, A. M., & Seifert, A. L. (2007). How do I feel about the behavior? The interplay of affective associations with behaviors and cognitive beliefs as influences on physical activity behavior. *Health Psychology, 26*, 152–158.
Kurdi, B., Gershman, S. J., & Banaji, M. R. (2019). Model-free and model-based learning process in the updating of explicit and implicit evaluations. *Proceedings of the National Academy of Science of the United States of America, 116*, 6035–6044.
Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron, 81*, 687–699.
Maher, J. P., & Conroy, D. E. (2015). Habit strength moderates the effects of daily action planning prompts on physical activity but not sedentary behavior. *Journal of Sport & Exercise Psychology, 37*, 97–107.
Markland, D., Hall, C. R., Duncan, L. R., & Simatovic, J. (2015). The effects of an imagery intervention on implicit and explicit exercise attitudes. *Psychology of Sport Exercise, 17*, 24–31.
Marteau, T. M., Hollands, G. J., & Fletcher, P. C. (2012). Changing human behavior to prevent disease: The importance of targeting automatic processes. *Science, 337*, 1492–1495.
McEachan, R. R. C., Conner, M., Taylor, N. J., & Lawton, R. J. (2011). Prospective prediction of health-related behaviors with the theory of planned behavior: A meta-analysis. *Health Psychology Review, 5*, 97–144.
Neal, D. T., Wood, W., Labrecque, J. S., & Phillippa, L. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *Journal of Experimental Social Psychology, 48*, 492–498.
Phillips, L. A., & Gardner, B. (2016). Habitual exercise instigation (vs. execution) predicts healthy adults' exercise frequency. *Health Psychology, 35*, 69–77.
Rebar, A. L., Dimmock, J. A., Jackson, B., Rhodes, R. E., Kates, A., Starling, J., & Vandelanotte, C. (2016). A systematic review of the effects of non-conscious regulatory process in physical activity. *Health Psychology Review, 10*, 395–407.
Rebar, A. L., Elavsky, S., Maher, J. P., Doerksen, S. E., & Conroy, D. E. (2014). Habits predict physical activity on days when intentions are weak. *Journal of Sport & Exercise Psychology, 36*, 157–165.
Rhodes, R. E. (2014). Adding depth to the next generation of physical activity models. *Exercise & Sport Sciences Reviews, 42*, 43–44.

Rhodes, R. E., Janssen, I., Bredin, S. S. D., Warburton, D. E. R., & Bauman, A. (2017). Physical activity: Health impact, prevalence, correlates and interventions. *Psychology & Health, 32*, 942–975.

Rhodes, R. E., Fiala, B., & Conner, M. (2009). A review and meta-analysis of affective judgments and physical activity in adult populations. *Annals of Behavioral Medicine, 38*, 180–204.

Rhodes, R. E., & Rebar, A. L. (2018). Physical activity habit: Complexities and controversies. In Verplanken, B. (Eds.), *The psychology of habit: Theory, mechanisms, change, and contexts* (pp. 91–109). Gewerbestrasse, Switzerland: Springer Nature Switzerland AG.

Rhodes, R. E., Saelens, B. E., & Sauvage-Mar, C. (2018). Understanding physical activity through interactions between the built environment and social cognition: A systematic review. *Sports Medicine, 48*, 1893–1912.

Sala, M., Baldwin, A. S., & Williams, D. M. (2016). Affective and cognitive predictors of affective response to exercise: Examining unique and overlapping variance. *Psychology of Sport & Exercise, 27*, 1–8.

Sniehotta, F. F., Presseau, J., & Araujo-Soares, V. (2014). Time to retire the theory of planned behavior. *Health Psychology Review, 8*, 1–7.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence, 112*, 181–211.

Tappe, K. A., & Glanz, K. (2013). Measurement of exercise habits and prediction of leisure-time activity in established exercise. *Psychology, Health & Medicine, 18*, 601–611.

Verplanken, B., Aarts, H., & Van Knippenberg, A. (1997). Habit, information acquisition, and the process of making travel mode choices. *European Journal of Social Psychology, 27*, 539–560.

Verplanken, B., & Melkevik, O. (2008). Predicting habit: The case of physical exercise. *Psychology of Sport & Exercise, 9*, 15–26.

Williams, D. M., & Evans, D. R. (2014). Current emotion research in health behavior science. *Emotion Review, 6*, 277–287.

Wood, W., & Runger, D. (2016). Psychology of habit. *Annual Reviews of Psychology, 67*, 289–314.

Dual system theory of physical activity: A reinforcement learning perspective

CHU Xin-Yu[1,3]; WANG Ze-Jun[2]; XIAO Huan-Yu[3]

([1] *Shanghai University of Engineering Science, Department of Physical Education, Shanghai 201620, China*) ([2] *Tongji University, International College of Football, Shanghai 200092, China*) ([3] *Shanghai University of Sport, School of Leisure Sport, Shanghai 200438, China*)

Abstract: Exercise behavior theory based on rational decision making is regarded as the dominant framework for understanding physical activity, which provides valuable information on

cognitive constructs linked to physical activity. Behavioral interventions based on social ecological model have attracted much attention due to its better performance. However, recent studies show that exercise benefits cognition and current sport environment do not promote the formation of individual exercise habit, so it is necessary to explore a new theoretical framework to clarify the formation mechanism of individual exercise habit. The latest framework for explaining physical activity is dual system model, which promises to provide a broader perspective on motivation by considering the non-conscious and hedonic determinants of physical activity. Multiple representative dual system models associated with physical activity, on the one hand, from simple spontaneous path to context cues and exercise habit, and then to complex concept model highlighting the role of automatic affective evaluations, the construct of system 1 is elaborated. Combined with system 2, which is concerned by exercise behavior theory, they offer a basis on how to build dual system model for physical activity. On the other hand, through analysis of the principle for competitive, cooperative and hierarchical control of dual systems, proposals are provided for the control of dual system model. Canonical reinforcement learning framework explains the principle for construct and control of dual system model: in terms of construct of the model, model-free and model-based reinforcement learning represent system 1 and system 2, respectively. In terms of control of the model, Dyna cooperative architecture and hierarchical reinforcement learning provide a reasonable explanation for physical activity is more likely to be a complex combination of actions which is hierarchically operated with a cooperative form than competitive control. Finally, the exerciser-sport environment interaction in reinforcement learning is put forward to explore exercise behavior from a brand-new perspective.
Keywords: physical activity; dual system model; theory of planned behavior; exercise habit; reinforcement learning